THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 273, No. 49, Issue of December 4, pp. 32685–32689, 1998
Printed in U.S.A.

Structural Basis for Inactivating Mutations and pH-dependent Activity of Avian Sarcoma Virus Integrase*

(Received for publication, August 28, 1998, and in revised form, September 21, 1998)

Jacek Lubkowski‡, Fan Yang‡, Jerry Alexandratos‡, George Merkel§, Richard A. Katz§, Kelly Gravuer§, Anna Marie Skalka§, and Alexander Wlodawer‡¶

From the ‡Macromolecular Structure Laboratory, ABL Basic Research Program, NCI-Frederick Cancer Research and Development Center, National Institutes of Health, Frederick, Maryland 21702 and the §Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Crystallographic studies of the catalytic core domain of avian sarcoma virus integrase (ASV IN) have provided the most detailed picture so far of the active site of this enzyme, which belongs to an important class of targets for designing drugs against AIDS. Recently, crystals of an inactive D64N mutant were obtained under conditions identical to those used for the native enzyme. Data were collected at different pH values and in the presence of divalent cations. Data were also collected at low pH for the crystals of the native ASV IN core domain. In the structures of native ASV IN at pH 6.0 and below, as well as in all structures of the D64N mutants, the side chain of the active site residue Asx-64 (Asx denotes Asn or Asp) is rotated by ~150° around the Cα—Cβ bond, compared with the structures at higher pH. In the new structures, this residue makes hydrogen bonds with the amide group of Asn-160, and thus, the usual metal-binding site, consisting of Asp-64, Asp-121, and Glu-157, is disrupted. Surprisingly, however, a single Zn^{2+} can still bind to Asp-121 in the mutant, without restoration of the activity of the enzyme. These structures have elucidated an unexpected mechanism of inactivation of the enzyme by lowering the pH or by mutation, in which a protonated side chain of Asx-64 changes its orientation and interaction partner.

Integrase (IN)[1] (1) is one of only four enzymes encoded by retroviruses, such as human immunodeficiency virus type 1 and avian sarcoma virus (ASV), and it is absolutely essential for the support of the viral life cycle. For these reasons, IN is currently a target for the design of antiretroviral drugs. Retroviral INs contain approximately 300 amino acids, organized into three domains: zinc-binding N terminus, catalytic core domain, and DNA-binding C terminus. IN catalyzes the incorporation of the reverse-transcribed viral DNA into the host genome in two steps, processing and joining (1–3), both of which involve a nucleophilic attack by a hydroxyl group on a DNA phosphate. In the processing step, a water molecule attacks near the end of the viral DNA, two nucleotides from the 3′-ends of both viral DNA strands. In the joining step, the exposed viral DNA deoxyribose 3′-OH is activated to attack the host DNA at a relatively nonspecific location, thereby inserting viral DNA into the host genome. *In vitro,* these reactions require only virus-like DNA, IN, and metal cations.

Although the isolated ASV IN catalytic core domain is defective for the processing and joining activities, it retains two activities: disintegration, the reverse of the joining step that uses a preformed DNA substrate, and an endonuclease activity that cleaves between the highly conserved C and A (−3 activity) at the termini of the viral DNA (CATT-3′) (4). The endonuclease activity is distinct from the normal processing activity of IN, which cleaves between the A and T (−2 activity). Both the endonuclease and disintegration activities of the catalytic domain are dependent on the D, D(35)E catalytic triad, corresponding to Asp-64, Asp-121, and Glu-157 in ASV IN, and an exogenous metal cation (5). The biological relevance of the 3′-endonuclease and disintegration activities is unclear; however, both activities are useful for monitoring the function of the active site.

In previous studies of the D, D(35)E motif, it was shown that conservative active site substitutions, such as Asp → Glu or Glu → Asp, reduced activity by 10-fold or more, and other substitutions, such as Asp → Ala, abolished it completely (5). These residues were therefore proposed to directly bind the required metal cofactors. Later studies showed that divalent cations, such as Ca^{2+}, Cd^{2+}, Zn^{2+}, Mg^{2+}, and Mn^{2+}, are tightly bound by Asp-64, Asp-121, or Glu-157 in the ASV IN active site, although only the last two cations support full catalytic activity (6, 7). The central residue of the metal-binding triad is Asp-64, which together with Asp-121 forms metal-binding site I, and with Glu-157 forms binding site II. It is clear from the nearly perfect coordination distances of metal cations (2.0–2.2 Å for O ... Mn^{2+}) that any substitution of the active site residues would disrupt metal binding by displacing the acidic oxygens.

Binding studies of human immunodeficiency virus type 1 IN and ASV IN with an inhibitor targeted against human immunodeficiency virus type 1 IN showed comparable inhibition effects, but the crystal structure of an inhibitor-catalytic core domain complex was determined only for ASV IN (8). One interesting observation during that study was the rotation of the side chain of the central acidic residue, Asp-64, around the Cα—Cβ bond. We hypothesized that this change resulted not directly from inhibitor binding, but rather from the decrease of the pH of the mother liquor caused by addition of the inhibitor. This interpretation was supported by the fact that although the inhibitor also bound to IN at a higher pH, a similar rotation of

* This research was sponsored in part by the NCI, National Institutes of Health, Department of Health and Human Services, under contract with ABL. Other support includes National Institutes of Health Grants CA-47486 and CA-06927, a grant for infectious disease research from Bristol-Myers Squibb Foundation, and an appropriation from the Commonwealth of Pennsylvania. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 1VSK, 1VSL, and 1VSM) have been deposited with the Protein Data Bank, Brookhaven National Laboratory, Upton, NY.

¶ To whom correspondence should be addressed. Tel.: 301-846-5036; Fax: 301-846-6128; E-mail: wlodawer@ncifcrf.gov.

[1] The abbreviations used are: IN, integrase; ASV, avian sarcoma virus; CHES, 2-(cyclohexylamino)ethanesulfonic acid; MES, 2-(morpholino)ethanesulfonic acid; MOPS, 3-(morpholino)propanesulfonic acid.

TABLE I
Key to the identification of the crystals of ASV IN catalytic core domain

Key	Definition of protein	pH of crystal growth or soaking solution used prior the data collection (buffer identity and concentration)	Identity and concentration of metal cation
ASVIN64	(Core) mutant D64N	6.0, citrate buffer, 0.1 M	None
ASVIN64Mg	(Core) mutant D64N	6.0, acetate buffer, 0.1 M	$MgCl_2$, 0.1 M
ASVIN64Mn	(Core) mutant D64N	6.0, acetate buffer, 0.1 M	$MnCl_2$, 0.1 M
ASVIN64Zn	(Core) mutant D64N	6.0, acetate buffer, 0.1 M	$ZnCl_2$, 0.1 M
ASVIN5	ASVIN core	5.0, acetate or citrate buffer, 0.1 M	None

TABLE II
Data collection and refinement statistics

	ASVIN64	ASVIN64Mg	ASVIN64Mn	ASVIN64Zn	ASVIN5
Data collection statistics					
Total number of reflections	85,055	46,546	59,328	33,008	28,268
Number of unique reflections	10,132	9450	9110	9425	9166
Resolution range (Å)	20.0–2.15	20.0–2.20	20.0–2.25	20.0–2.20	20.0–2.13
Completeness of the data (%)	97.0	97.3	99.9	96.8	85.7
Highest resolution shell (limit, Å)	96.5 (2.21)	98.9 (2.26)	99.9 (2.32)	98.0 (2.26)	86.5 (2.19)
R_{sym}	0.077	0.064	0.083	0.068	0.080
Average I/σ(I)	20.7	16.4	16.2	12.8	12.8
In the last shell	3.6	4.0	4.0	2.2	2.3
Refinement statistics[a]					
Unit cell parameters (a/c, Å)	67.16/79.65	67.01/79.71	67.04/79.78	66.97/79.99	67.09/79.24
Resolution ranges (Å)	8.0–2.20	8.0–2.20	8.0–2.25	8.0–2.20	8.0–2.15
Crystallographic R-factor	15.9	15.1	14.9	15.6	14.6
Free R-factor	24.9	22.6	23.4	24.4	23.8
rmsd[b] from ideality-Bonds (Å)	0.014	0.013	0.013	0.015	0.013
Angles (degrees)	1.84	1.70	1.75	1.75	1.76
Impropers (degrees)	2.28	2.41	2.39	2.39	2.35
Dihedrals (degrees)	25.5	25.5	24.6	25.6	24.4
Average B-factor (Å^2, all atoms)	34.8	33.6	32.5	34.4	32
Side chain of Asp/Asn-64	24.9	24.4	23.0	26.3	24.0
Side chain of Asp-121	40.3	36.5	37.4	39.4	37.9
Side chain of Glu-157	65.6	72.2	69.3	65.5	53.7
Number of non-hydrogen atoms	1230	1245	1238	1226	1249
Number of solvent atoms	106	121	114	100	125
Number of metal cations	0	0	0	2	0

[a] All structures were refined against x-ray data corresponding to $|F| \geq 2 \cdot \sigma(F)$.
[b] rmsd, root mean square deviation.

the side chain of Asp-64 was not observed. In the experiments carried out at lower pH, the structural transition was determined to be between 5.5 and 5.6, *i.e.* near the pK_a of the carboxylic acid group of the aspartate side chain. In this paper, we report the structural basis of the inactivation of the ASV IN catalytic core domain resulting from a single conservative active site mutation (D64N), as well as from the decrease of pH, and we correlate the inactivation with the ability of the enzyme to bind divalent cations.

MATERIALS AND METHODS

Activity Profile of the Full-length ASV IN and the Catalytic Core—The endonuclease activity of the catalytic core domain was assayed as described previously, using 10 mM $MnCl_2$ for 10 min at 37 °C (9). The following buffer systems were used: CHES, Tris, MOPS, MES, and acetate, providing a pH range of 4.3–9.8.

Preparation of Crystals—Crystals of both the catalytic core domain of ASV IN, and the core domain D64N mutant, ASVIN64 (definitions of the abbreviations for different crystals described in this report are given in Table I), were obtained under identical conditions, described previously (9). Unless otherwise stated, citrate buffer was used to adjust the pH of the crystallization medium to 6.0. Before the x-ray experiments, some ASV IN crystals were transferred stepwise to synthetic mother liquor solutions with gradually decreasing pH. The pH value of the final soaking solution for these crystals was 5.0. Therefore, "low pH" is used throughout this paper to refer to these crystals and the structures that originated from them. Whenever soaking was accompanied by the addition of divalent cations, acetate buffer was used instead of citrate buffer at an equivalent concentration and pH. In this manner, some of the ASVIN64 crystals were first transferred to acetate buffer (pH 6.0) and then soaked in the presence of Mn^{2+}, Zn^{2+}, or Mg^{2+} for several days. All soaking experiments were performed at 4 °C.

For the x-ray diffraction experiments with the low pH crystals, we selected those soaked in the absence of divalent cations (ASVIN5) and those soaked in the presence of 0.1 M $MnCl_2$. In the case of the D64N mutant, data were collected for crystals not exposed to divalent cations, as well as for those soaked in the presence of 0.1 M $MnCl_2$, $MgCl_2$, or $ZnCl_2$.

Data Collection and Structure Refinement—Diffraction data for all crystals were collected at room temperature using CuKα radiation generated by a rotating-anode Nonius FR591 generator operated at 45 kV and 95 mA and were recorded on a DIP2020 image plate detector. All data sets were processed with the program package DENZO and scaled with SCALEPACK (10). Data collection statistics for the sets used in the studies are shown in Table II.

The refinement of all structures was carried out with the program X-PLOR (11). The structure of ASV IN deposited at the Protein Data Bank with accession code 1asv was used as the starting model. This structure corresponds to the crystals grown in the presence of citrate buffer (9); therefore, we considered it to be the most relevant to this study. Before the initial rigid body refinement was performed, we removed all solvent molecules from the model, as well as residues 144–152, which form a part of the flexible active site loop. The side chain of Asp-64 beyond Cβ was also removed. The refinement protocol was cross-validated by the free R-factor index (12), calculated for ~10% of all reflections, which were excluded from the refinement. Rigid body refinement at the resolution range 3.0–8.0 Å was carried out to compensate for small differences in the unit cell parameters. This step was followed by the positional and overall B-factor refinements. After visual inspection of the structures with the programs O (13) or Chain (14) and their manual corrections, the resolution of the data was extended and the B-factors were refined individually for all non-hydrogen atoms. At the final stages of the refinement, water molecules, as well as metal cations, if present, were positioned as indicated by appropriate peaks in the difference electron density maps. The missing fragments of the protein chain were also reconstructed, whenever possible, on the basis of the electron density (*e.g.* for the side chain of residue 64). The quality

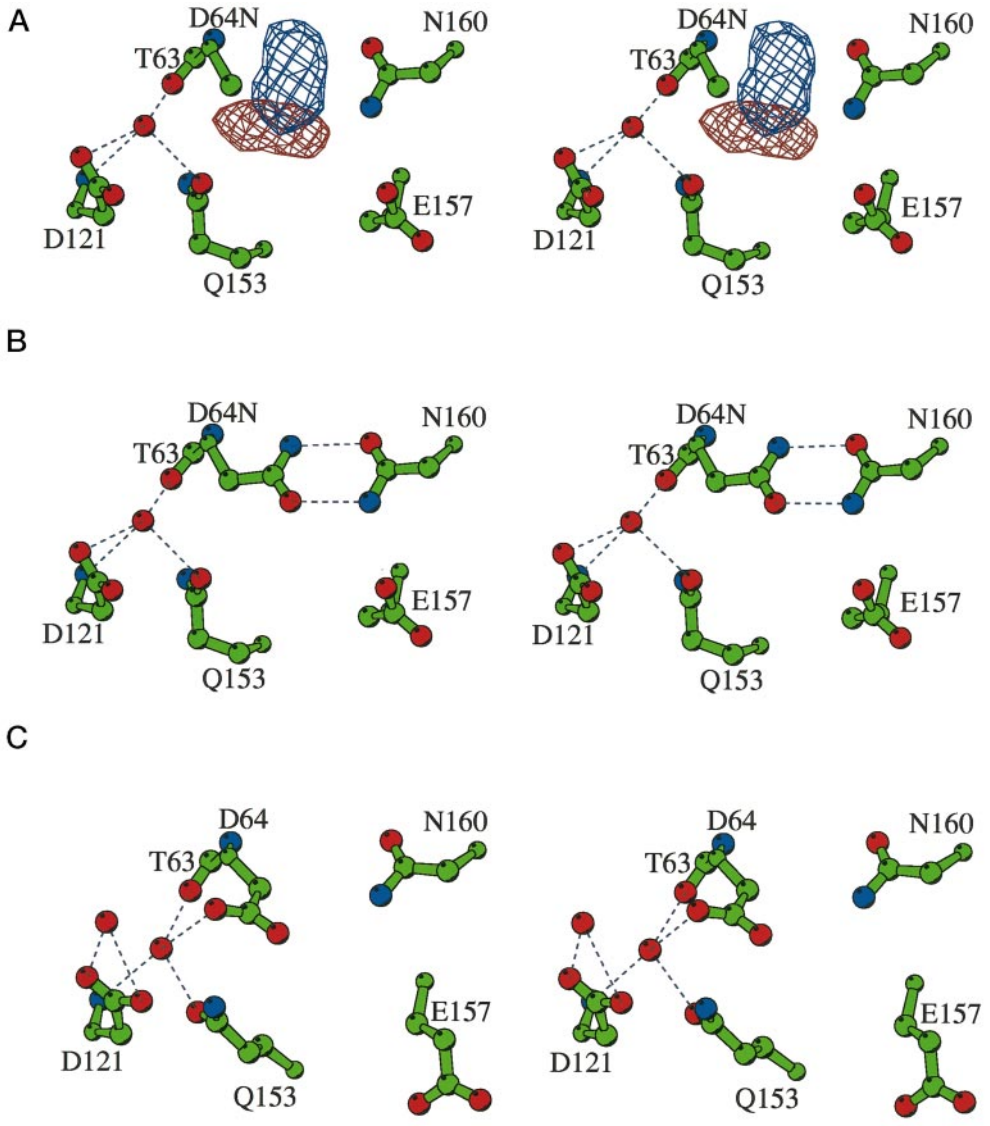

FIG. 1. **The active site of ASV IN (catalytic core domain).** *A*, $F_o - F_c$ difference density "omit" map, contoured at 2.5 σ, showing the position of the side chain of D64N (*blue*) compared with the side chain of Asp-64 (*brown*) in the active conformation. The hydrogen bonding pattern of a structurally conserved water molecule observed in all ASV IN structures is also shown. *B*, hydrogen bonding of the side chain of D64N with Asn-160. The hydrogen bonding pattern of a structurally conserved water molecule observed in all ASV IN structures is also shown. *C*, hydrogen bonding of the side chain of Asp-64 with the structurally conserved water.

of the geometrical and stereochemical indices was continuously monitored during refinement with the program PROCHECK (15). The refinement statistics and quality assessments are shown in Table II.

RESULTS

Conformation of Asp/Asn-64—In all the structures discussed here (except that of the ASV IN catalytic core soaked at pH 5.0 in the presence of Mn^{2+}), we observed a similar conformation of the side chain of residue 64, which was different from the one reported previously (5). In the previously reported structures of ASV IN (6, 7, 9), with the sole exception of the low pH complexes with an inhibitor (8), the side chain of Asp-64 points toward the side chain of Asp-121, forming hydrogen bonds mediated by a water molecule (or a metal ion). The difference $F_o - F_c$ electron density extending from Cβ of residue 64 unambiguously indicates that the Asx-64 side chain (Asn in the mutant ASVIN64 or Asp in ASVIN5) is rotated by ~150° around the Cα—Cβ bond (Fig. 1*A*), compared with the standard orientation (Fig. 1*B*). In the structures reported here, Asx-64 forms two strong hydrogen bonds with the side chain of Asn-160. In the case of the mutant, introduction of the amide group to the side chain of residue 64 opens the possibility of the formation of two hydrogen bonds (N-H . . . O distances of 2.8–3.0 Å) with the amide group of Asn-160. Because the charged side chain of Asp-64 could not participate in equivalent hydrogen bonds, its conformation in the structure of ASVIN5 can be explained by the predominant presence of the protonated form under these experimental conditions. In the structure of the low pH ASV IN derived from the crystals soaked in the presence of Mn^{2+}, the difference electron density indicates the native orientation (Fig. 1*C*). Moreover, for this structure we can also identify an electron density peak corresponding to the Mn^{2+} bound to the primary metal-binding site of ASV IN (6). However, the presence of divalent metal cations (Mn^{2+}, Mg^{2+}, and Zn^{2+}) at 0.1 M concentration in the solutions did not alter the conformation of the Asn-64 side chain in the mutant (Fig. 1*A*).

Binding of the Divalent Cations—As described above, ASV IN was capable of binding Mn^{2+} even at low pH, by repositioning of the Asp-64 side chain to the native orientation. Similar experiments performed with crystals of the mutant D64N (crystals ASVIN64Mn and ASVIN64Mg) did not indicate any binding of either Mg^{2+} or Mn^{2+} (data not shown). In the case of

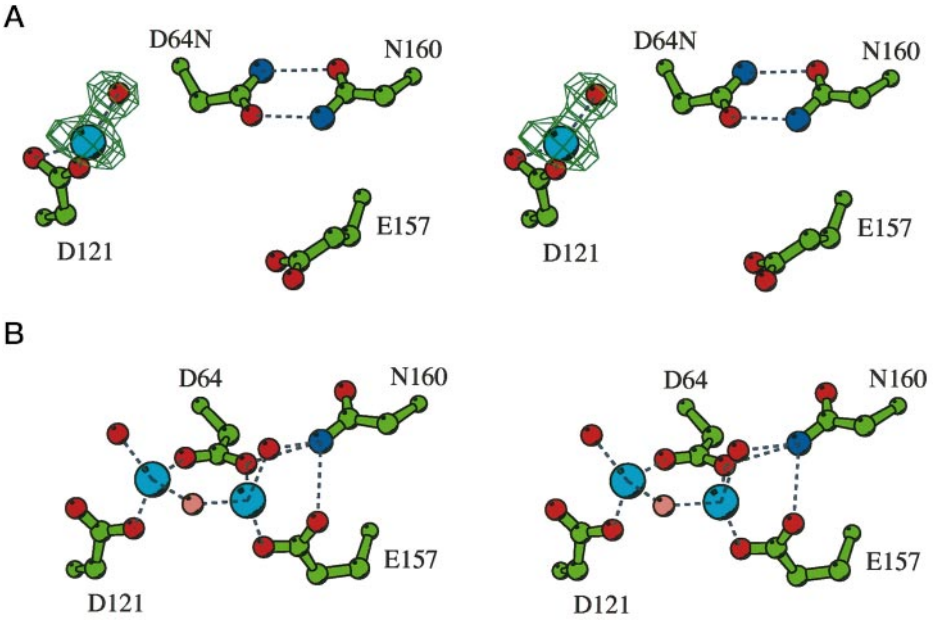

FIG. 2. **The active site of ASV IN (catalytic core domain) with bound zinc cation(s).** *A*, $2F_o - F_c$ density map, contoured at 2.5 σ, showing the bound Zn^{2+} hydrogen bonded only to the side chain of Asp-121. The side chain of D64N is hydrogen bonded to Asn-160. *B*, a previously determined structure of ASV IN, showing the bound zinc cations hydrogen bonded to the side chains of Asp-64 and Asp-121.

crystals soaked in 0.1 M $ZnCl_2$ (ASVIN64Zn), we observed a single, very clear peak in the positive $F_o - F_c$ density at 2.07 Å from an OD atom of Asp-121 (Fig. 2*A*). This peak could be interpreted only as a Zn^{2+} bound to the single active site residue. We did not observe the second Zn^{2+} previously reported bound between the side chains of Asp-64 and Glu-157 (Fig. 2*B*) (7). Two additional zinc-binding sites, located away from the active site, were also previously reported in the structure of ASV IN with Zn^{2+}. In the case of D64N IN, we also observed Zn^{2+} in close proximity to His-103, corresponding to site III reported for the native ASV IN. We were unable to locate another Zn^{2+} at site IV between His-198 and Tyr-194 (7).

Conformation of the Flexible Active Site Loop—In none of the structures reported here were we able to trace completely the active site loop (residues 144–154), due to the poor quality of the electron density in this area. In some structures (ASVIN64 and ASVIN5), interpretable electron density covered residues 144–146 and 153–154. The overall trace was quite apparent, although the quality of the electron density indicated the highly dynamic character of these residues; therefore, their precise conformation could be questioned. Similarly to our previous observations (8), the tracing of the visible part of this loop is different from the conformations described in our first report (9). A detailed analysis of the loop conformation cannot be completed with the currently available data.

Effect of pH on Activity of the Catalytic Core—Having observed pH-dependent differences in the conformation of the Asp-64 side chain, we next investigated the effect of pH on the endonuclease activity of the catalytic core domain. We previously reported that the optimal pH for the full-length wild type ASV IN was ~9 (16). Using the IN catalytic core domain as well as the full-length wild type IN protein, we carried out a standard assay for IN processing and endonuclease activity. The results show that the optimal pH for the processing activity of full-length IN is ~8.5, which is generally consistent with the earlier results mentioned above (Fig. 3). The −3′ processing activity of the catalytic core also showed an optimal pH of ~8.5. At pH 5, both proteins were essentially inactive. We therefore conclude that the pH activity profile is consistent with the observation that the Asp-64 side chain is repositioned at pH 5.

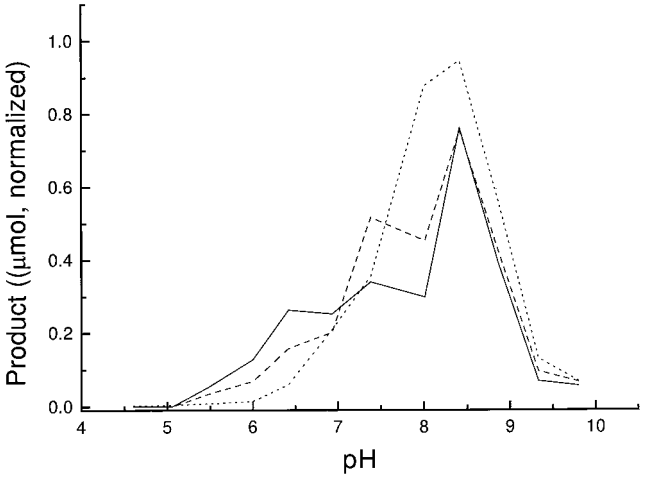

FIG. 3. **pH dependence of the endonuclease activity of ASV IN.** The reaction was performed as noted under "Materials and Methods." The −2 and −3 activities of the full-length protein are indicated by *solid* and *dashed lines*, respectively. The −3 endonucleolytic activity of the catalytic core is indicated by a *dotted line*.

DISCUSSION

The main goal of the studies described here was to understand the molecular basis for the mutations that inactivate the ASV IN catalytic core. The problem was approached via two series of experiments, namely by studies of the D64N mutant, and of the wild type core domain at a pH significantly lower than physiological. As shown previously (5), mutation of Asp-64 to Ala essentially inactivates the enzyme. Several structures of the conservative D64N mutant, reported here, clearly show disruption of the primary metal-binding site resulting from the rotation of the side chain of Asn-64. Both observed conformations of this side chain appear to be energetically favorable. In the first mode of binding, predominantly observed in the previously reported structures (6, 7, 9), Asp-64 is anchored by a hydrogen bond to a water molecule, which in turn is stabilized by an additional hydrogen bond with the side chain of Asp-121. Because it was shown previously that this conformation of Asp-64 remains unchanged upon binding of several divalent metal cations, we refer to it as the "active" conformation. The ability of ASV IN to bind divalent cations with high affinity suggests that the side chains of both Asp-64 and Asp-121 are most likely present in their ionized forms at neutral pH. Therefore, the presence of a water molecule that is hydrogen bonded with the two side chains seems to be crucial for the stability of

this conformation. In fact, in all of the structures of the ASV IN catalytic core reported previously, as well as in the structures presented here, this water molecule is completely structurally conserved, participating in hydrogen bonding interactions with several protein atoms. In addition to being involved in hydrogen bonds with the side chains of Asp-64 and Asp-121, this water molecule also serves as a proton donor in the hydrogen bond with Thr-63(O) and as a proton acceptor in the hydrogen bond with Asp-121(N). Furthermore, it interacts, most likely by donating a proton, with Asp-121(OD1), and with either OD1 or NH_2 of Gln-153.

In the second conformation, observed for the D64N mutant, the side chain of Asn-64 interacts directly with the side chain of Asn-160. For the wild type ASV IN, the equivalent interaction may depend upon the protonation state of Asp-64. With a protonated aspartate, such a conformation would probably be quite favorable, as it would lead to formation of two strong hydrogen bonds, similar to the ones present in the D64N mutant. For an ionized aspartate, however, only one hydrogen bond could be formed, via protons of the amide group of Asn-160. Moreover, the carboxylate oxygen atom of the side chain of Asn-160 would be in a repulsive interaction with an oxygen of the carboxylate of Asp-64. The pH 5.0 structure is consistent with such a description, with the conformation of the side chain of Asp-64 identical to that found in the D64N mutant. On the basis of the current results we cannot postulate a unique $C\beta$—$C\gamma$ rotamer of the side chain of residue 64. The network of hydrogen bonds around the side chain of Asn-160 does not favor a specific rotamer in any of the available structures. Therefore, it is possible that within the pair of residues Asx-64 and Asn-160, the side chains are 2-fold disordered. A detailed description of the conformation of residue 64 might be a very important element in the understanding of the activity of this enzyme on a molecular basis. This is true because only the active conformation of ionized Asp-64 allows normal binding of divalent cations.

Additional observations are related to the conformation of the flexible loop adjacent to the active site, comprising residues 144–154. This is the second time we noticed a loop conformation different from those described before (9), as previously reported for the structure of a complex between ASV IN and an inhibitor (8). Two principal factors, the orientation of residue 64 and the pH of the mother liquor in which crystals were grown or soaked, seem to be associated with the loop conformation. It is possible that residue 64 in its inactive conformation (such as in the mutant structure presented here) allows stabilization of the active site loop in a specific conformation. Furthermore, the protonation state of some loop residues or residues interacting with the loop might change when the pH drops, forming an environment that is suitable for the stabilization of a specific conformation. Recently, we collected ultra-high resolution x-ray data for ASV IN crystals grown under different conditions, and the preliminary results seem to support the possibilities described above. A conclusive analysis will be possible only after these new structures are completely refined.

REFERENCES

1. Katz, R. A., and Skalka, A. M. (1994) *Annu. Rev. Biochem.* **63,** 133–173
2. Vink, C., and Plasterk, R. H. (1993) *Trends Genet.* **9,** 433–438
3. Goff, S. P. (1992) *Annu. Rev. Genet.* **26,** 527–544
4. Kulkosky, J., Katz, R. A., Merkel, G., and Skalka, A. M. (1995) *Virology* **206,** 448–456
5. Kulkosky, J., Jones, K. S., Katz, R. A., Mack, J. P., and Skalka, A. M. (1992) *Mol. Cell. Biol.* **12,** 2331–2338
6. Bujacz, G., Jaskólski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R. A., and Skalka, A. M. (1996) *Structure* **4,** 89–96
7. Bujacz, G., Alexandratos, J., Wlodawer, A., Merkel, G., Andrake, M., Katz, R. A., and Skalka, A. M. (1997) *J. Biol. Chem.* **272,** 18161–18168
8. Lubkowski, J., Yang, F., Alexandratos, J., Wlodawer, A., Zhao, H., Burke, T. R., Jr., Neamati, N., Pommier, Y., Merkel, G., and Skalka, A. M. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 4831–4836
9. Bujacz, G., Jaskólski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R. A., and Skalka, A. M. (1995) *J. Mol. Biol.* **253,** 333–346
10. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
11. Brünger, A. (1992) *X-PLOR,* Version 3.1, Yale University Press, New Haven, CT
12. Brünger, A. T. (1992) *Nature* **355,** 472–474
13. Jones, T. A., and Kieldgaard, M. (1997) *Methods Enzymol.* **277,** 173–208
14. Sack, J. S. (1988) *J. Mol. Graph.* **6,** 244–245
15. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
16. Muller, B., Jones, K. S., Merkel, G. W., and Skalka, A. M. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 11633–11637